# Découverte de l'Intelligence Artificielle

*Examen sur table : Option L1 et CPES*

Simon Collet, Thomas Deneux, George Marchment

**Date :** 21/12/2023

Numéro de copie anonymisée : [ ]

## Exercice 1 : Questions de cours

1. Définissez chacun des termes suivants en quelques lignes chacun.

- Modèle (dans le contexte du Machine Learning)

Un modèle en Machine Learning est une **fonction** (donc qui prend des entrées et renvoie des sorties) qui dépend de **paramètres** (ce sont ces paramètres qui vont être modifiés lors de l'apprentissage).

- Apprentissage supervisé

L'apprentissage supervisé est un des 3 grands modes d'apprentissage en Machine Learning. Il consiste à entraîner le modèle à partir de données consistant en des paires (entrée, sortie), aussi appelées communément **(données, étiquettes)** dans le cas d'un modèle de classification. **L'entraînement consiste à faire en sorte que les sorties du modèle sur les données d'entrée d'entraînement coïncident avec les sorties correspondantes.**

- Apprentissage non supervisé

L'apprentissage non supervisé est un des 3 grands modes d'apprentissage en Machine Learning. Il consiste à entraîner le modèle à partir **uniquement de données d'entrées**, dont les **structures vont être détectées automatiquement**, sans étiquetage par un

opérateur humain. Les sorties du modèle pourront consister en des catégories (on parle alors de *clustering*) ou des décompositions des données d'entrée en sous-composantes (on parle de *réduction de dimension*).

- Apprentissage par renforcement

L'apprentissage par renforcement est un des 3 grands modes d'apprentissage en Machine Learning. Dans ce cas le modèle est appelé un **agent**, et interagit avec **environnement** : l'agent peut être par exemple un robot dans un environnement réel, ou un joueur virtuel dans un jeu de société (ex. jeu de Go) ou un jeu vidéo. Il reçoit de l'environnement l'information sur son **état** et doit décider d'actions à effectuer, et suite auxquelles il reçoit un nouvel état, et une **récompense**. Le but de l'apprentissage sera d'apprendre à choisir les **actions** permettant d'obtenir le maximum de récompenses.

2. Pour chacun des trois algorithmes suivants, précisez s'ils relèvent de l'apprentissage supervisé, non supervisé ou par renforcement et expliquez leur principe en une courte phrase (sans entrer dans aucun détail mathématique).

- KNN

**Apprentissage supervisé**.
La prise de décision pour un nouveau point est basée sur la **recherche dans la base de données d'entraînement du ou des points à la distance la plus petite** du point considéré : la décision prise est alors la **décision majoritaire** au sein de ces « plus proches voisins ».

- Réseau de neurones artificiels

**Apprentisssage supervisé.**
Un réseau de neurones (aussi appelé apprentissage profond ou « deep learning » en anglais) a une architecture imitant le fonctionnement du cerveau biologique, étant organisé en **couches** successives de **neurones artificiels**. Les neurones des couches intermédiaires et de sortie ont chacun leur activité calculée comme la **somme pondérée** des activités des neurones de la couche précédente, passée ensuite à une *fonction non-linéaire dite d'activation*.

- Q-learning

**Apprentissage par renforcement**.
L'algorithme apprend les **valeurs** associées à chaque action possible *a* partant de chaque état possible *s*, cette valeur est généralement notée Q(s,a). Pour l'algorithme Q-learning l'ensemble des états possibles et celui des actions possibles doivent tous deux être *finis*, en effet, les valeurs Q(s,a) sont stockées simplement dans une **table**. La valeur Q(s,a) de l'action a depuis l'état s est **liée aux récompenses** qui seront potentiellement reçues après avoir effectué cette action (*récompense immédiate*) mais également après les actions qui suivront (*récompenses futures*, elles-mêmes approximées lors du calcul par $\max_{a'} Q(s',a')$, s' étant le nouvel état atteint et a' les possibles actions suivantes).

3. Qu'appelle-t-on les "biais des IA" ? Nous avons détaillé en cours deux causes distinctes de biais, dans le cadre de l'apprentissage supervisé. Expliquez ces deux causes en donnant à chaque fois un exemple dans le cas d'une tâche de reconnaissance d'image.

Les biais des IA sont les **erreurs** des modèles d'IA une fois en utilisation (c'est-à-dire après leur entraînement).
En cours nous avons distingué les « biais explicite » qui reproduisent des **erreurs déjà présentes dans la base d'entraînement** (exemple : le robot tourne dans la mauvaise direction parce que on l'a entraîné à aller dans cette mauvaise direction) et « biais implicite » qui sont des **erreurs de généralisation** à des données relativement différentes de celles présentes dans la base de données (exemple : le robot veut continuer d'avancer quand il est bloqué par un mur, cette situation n'ayant pas été vue pendant l'entraînement).

## Exercice 2 : Interprétation code Python

Analysez le code Python ci-dessous, et répondez aux questions.

```
def fonction_mystere(liste):
    indice = 0
    val = liste[0]
    for i in range(1, len(liste)):
        if liste[i] > val:
            indice = i
            val = liste[i]
    return indice

ma_liste = [12, 45, 67, 23, 56, 89, 34]
val_mystere = fonction_mystere(ma_liste)
```

1. Quelle tâche effectue la fonction “`fonction_mystere`” ?

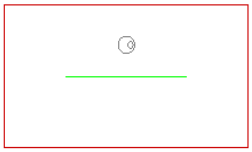
Elle renvoie l’indice du maximum dans une liste.

2. Que contient la variable “`val_mystere`” après exécution de ce code ?

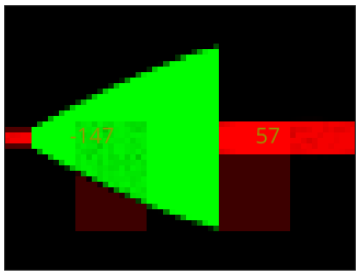
Elle renvoie 5. (Attention, les indices commençant à zéro, le sixième élément dans la liste a l’indice 5).

# Exercice 3 : Interprétation de l'espace d'états

Ci-dessous, nous reprenons la visualisation de l'espace des états d'entrée de l'IA comme dans le cas du dernier TP. Pour rappel, voici le plan de l'arène :

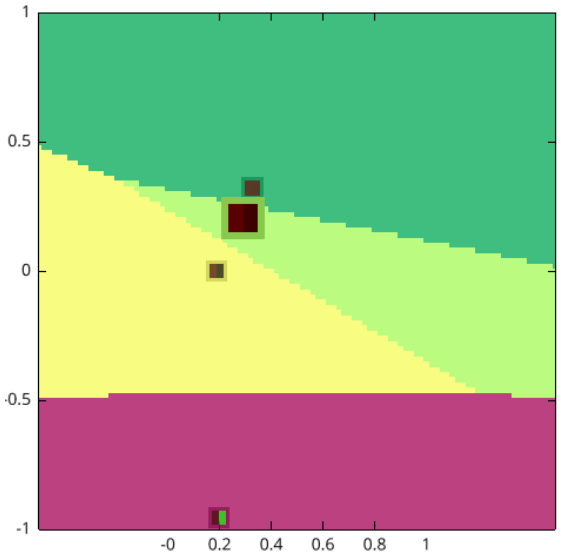


Notre robot simulé reçoit 2 entrées qui sont le calcul "canal rouge - canal vert" pour 2 zones de l'image à gauche et à droite de son champ visuel. Dans l'exemple ci-dessous, ces deux valeurs sont -147 et 57. La première valeur, négative, correspond à la présence d'un mur vert (mur intérieur). La seconde valeur, positive, correspond à un mur rouge (mur extérieur). Une valeur proche de zéro correspond à une quantité équivalente de rouge et de vert.

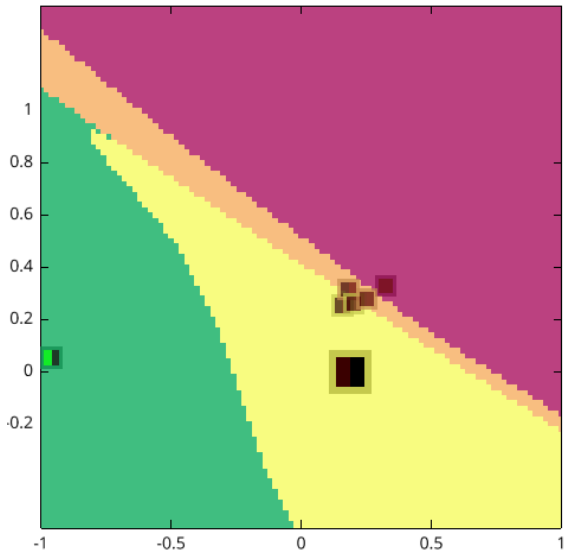



Les valeurs sont comprises entre -255 et 255 sur l'image de la caméra mais sont renormalisées entre -1 et 1 sur le graphe d'états.

Ci-dessous, nous représentons deux graphes d'états différents suite à l'entraînement de deux algorithmes distincts.




Entraînement n°1

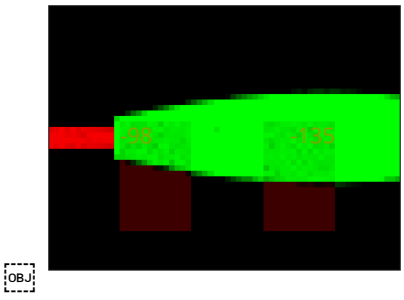



Entraînement n°2

Les couleurs des points et de l'arrière-plan correspondent aux actions disponibles au robot ci-dessous ; les points carrés sont les données d'entraînement (le point plus gros n'est pas une donnée d'entraînement, mais dans les deux cas il cache une donnée d'entraînement).

Répondez aux questions suivantes :

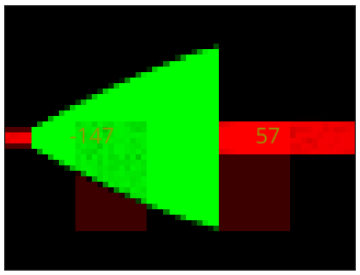
1. Si le robot se trouve dans la configuration suivante, quelle action va-t-il effectuer (répondez pour chacun des deux entraînements) ?




À gauche on lit "-98" et à droite "-135"

Entraînement n°1 : pivoter à gauche (couleur rose ; remarque : si on avait eu à droite une valeur un peu moins négative, par exemple -90, cela aurait été couleur jaune = tout droit). Entraînement n°2 : pivoter à droite.

2. Si le robot se trouve dans la configuration suivante, quelle action va-t-il effectuer (répondez pour les deux entraînements) ?




À gauche on lit “-147” et à droite “57”

Entraînement n°1 : tout droit. n°2 : pivoter à droite.

3. Quel entraînement est le plus performant pour que le robot se déplace autour de l’arène dans le sens d'une aiguille d'une montre (Justifier)?

Avec l’entraînement n°1 : le robot tourne à gauche lorsque les deux pixels voient du vert, et à droite lorsqu’ils voient du rouge ; il va tout droit lorsqu’ils voient du blanc ; le robot va donc s’orienter sur le circuit de telle sorte qu’il ait un mur vert à sa droite et un mur rouge à sa gauche, et tournera dans le sens des aiguilles d’une montre.

4. Quels algorithmes d’apprentissage ont été utilisés pour obtenir ces deux espaces d'états ? Quels sont les indices qui vous permettent de le déduire ?

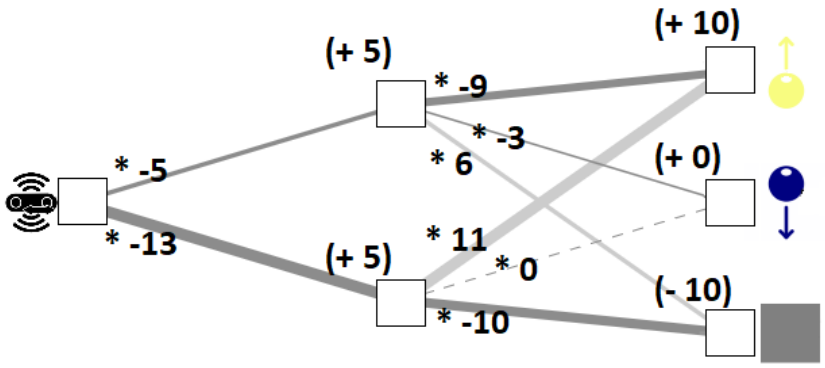
Entraînement n°1 : **KNN** avec K = 1. En effet on voit bien que les frontières sont les **droites médiatrices** entre les deux points les plus proches.

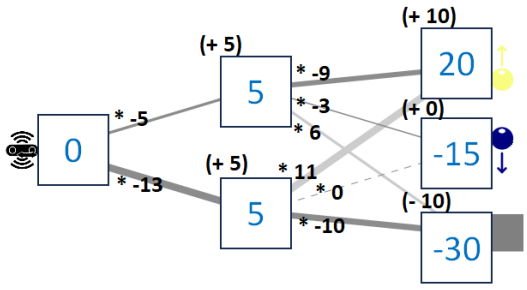
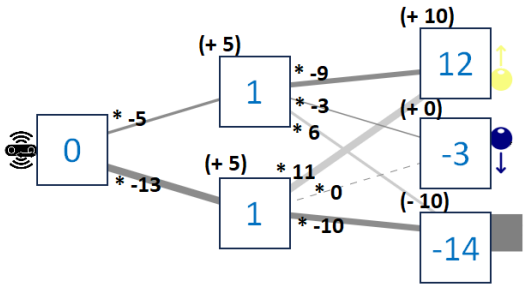
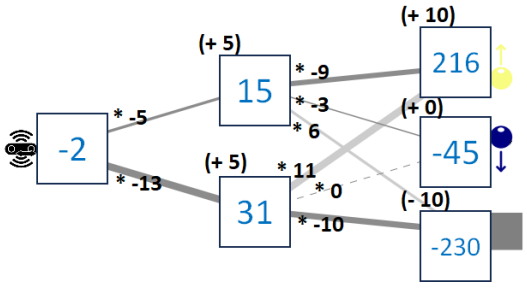
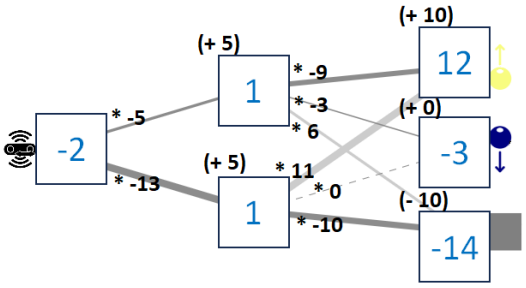
Entraînement n°2 : **réseau de neurones** avec **couche(s) intermédiaire(s)**. En effet les frontières n’étant **pas des droites**, cela ne peut pas être KNN, ni un réseau de neurones sans couche intermédiaire (et la fonction d’activation n’est pas affine par morceaux !).

# Exercice 4 : Réseau de neurones

Pour le réseau de neurones ci-dessous, calculez les activations des neurones de la couche intermédiaire et des neurones de sortie, avec les fonctions d'activation et différentes valeurs d'entrée ci-dessous. **Rappel :** la fonction d'activation va s'appliquer sur les deux neurones de la couche intermédiaire, mais pas sur la couche de sortie.

Attention, la fonction d'activation ne s'applique pas non plus à la couche d'entrée !!




| | Fonction d'activation **Leaky ReLU** $y = \begin{cases} \frac{z}{10} \; si \; z < 0 \\ z \; si \; z \geq 0 \end{cases}$ | Fonction d'activation **Heaviside** $y = \begin{cases} 0 \; si \; z < 0 \\ 1 \; si \; z \geq 0 \end{cases}$ |
|---|---|---|
| Valeur d'input : 0 |   |   |
| Valeur d'input : -2 |   |   |

# Exercice 5 : Apprentissage du morpion en Q-learning

Nous allons utiliser le Q-learning pour trouver la stratégie optimale pour jouer au morpion contre un joueur “moyen”. Mais commençons par quelques questions de cours.

**Question 1 : à propos de la formule ci-contre.**

$$target = (1-\gamma)r + \gamma \max_a Q(s', a)$$

Comment appelle-t-on le paramètre $\gamma$ ?
Facteur d’actualisation

Quelles sont ses bornes (valeurs minimales et maximales) ?
0 et 1

**Question 2 : à propos de la formule ci-contre.**

$$Q'(s,a) = (1-\alpha)Q(s,a) + \alpha\ target$$

Comment appelle-t-on le paramètre $\alpha$ ?
Vitesse d’apprentissage

Quelles sont ses bornes (valeurs minimales et maximales) ?
0 et 1

Le jeu du morpion (tic-tac-toe) est un jeu à deux joueurs qui se pratique sur une grille de taille 3x3. Les joueurs placent chacun à leur tour un symbole dans une case inoccupée. Le jeu se termine lorsqu’un joueur parvient à aligner 3 de ses symboles en ligne ou en diagonale, ce joueur est alors vainqueur ; ou lorsque toute les cases du plateau sont occupées sans qu’un joueur ait réussi à aligner 3 symboles : le résultat est alors un match nul.

La figure ci-dessous (page 10) est un arbre représentant différentes possibilités de déroulement d’une partie de morpion (en se restreignant aux coups les plus pertinents, sinon l’arbre de tous les coups possibles serait trop gros). Le premier joueur sera l’IA, elle utilise le symbole X ; le second joueur utilise le symbole O. Nous allons supposer que ce

second joueur (symbole O) joue de manière aléatoire entre les coups représentés dans l'arbre : à partir d'une position donnée, il choisit de manière équiprobable un coup parmi les possibilités représentées. Les probabilités de ces choix ont été représentées en vert sur la figure.

L'objectif de l'exercice est de déterminer la meilleure stratégie pour l'IA grâce à l'algorithme du Q-learning. Nous allons donc fixer une récompense de 100 en cas de victoire de la joueuse X, une récompense négative de -100 en cas de victoire du joueur O, et une récompense nulle en cas de match nul. **Attention :** la récompense n'est gagnée que lorsque le jeu est terminé, les coups joués en cours de partie ne rapportent pas de récompense immédiate !

Plutôt que d'appliquer l'algorithme du Q-learning étape par étape, nous pourrons déterminer directement vers quelles valeurs limites l'algorithme va converger, en utilisant l'équation de Bellman. En effet, au terme de l'apprentissage les Q-valeurs vérifient l'équation de Bellman : pour tout état $s$ et action $a$, la Q-valeur $Q(s,a)$ vérifie :

$$Q(s,a) = E[r + \gamma \, max_{a'} \, Q(s',a')],$$

où $E$ représente l'espérance mathématique (c'est à dire la moyenne en probabilité), $s'$ le nouvel état atteint et $max_{a'} \, Q(s',a')$ la meilleure Q-valeur possible à partir de ce nouvel état.

Nous utiliserons $\gamma$ = 1, ce qui est le plus approprié dans la résolution d'un jeu.

**Question 3 :**

Que représente l'*état s* dans le cadre du jeu du morpion ?
L'état de la grille, avec les positions des symboles X et O

**Question 4 :**

Que représente l'*action a* dans le cadre du jeu du morpion ?
Un coup d'un joueur, c'est-à-dire placer un symbole X ou O dans une case

**Question 5 :**

Que représente la notion d'*environnement* en apprentissage par renforcement ? Et en quoi consiste l'environnement ici dans le cadre du jeu du morpion ?

L'environnement représente l'environnement à l'intérieur duquel l'agent évolue. Plus spécifiquement, l'environnement envoie initialement à l'agent sont état initial, puis à chaque fois que l'agent effectue une action lui **renvoie une valeur de récompense et son nouvel état**.
Ici l'environnement

**Question 6 :**

Dans la figure page 10, nous avons complété les Q-valeurs de trois couples (s, a), à droite et vers le haut du graphe :

- dans un des trois cas, le coup donne lieu de manière univoque à un nul, c'est à dire récompense 0 ; on a donc

  $$Q(s,a) = E[r + \gamma\, max_{a'}\, Q(s',a')] = E[r] = 0$$

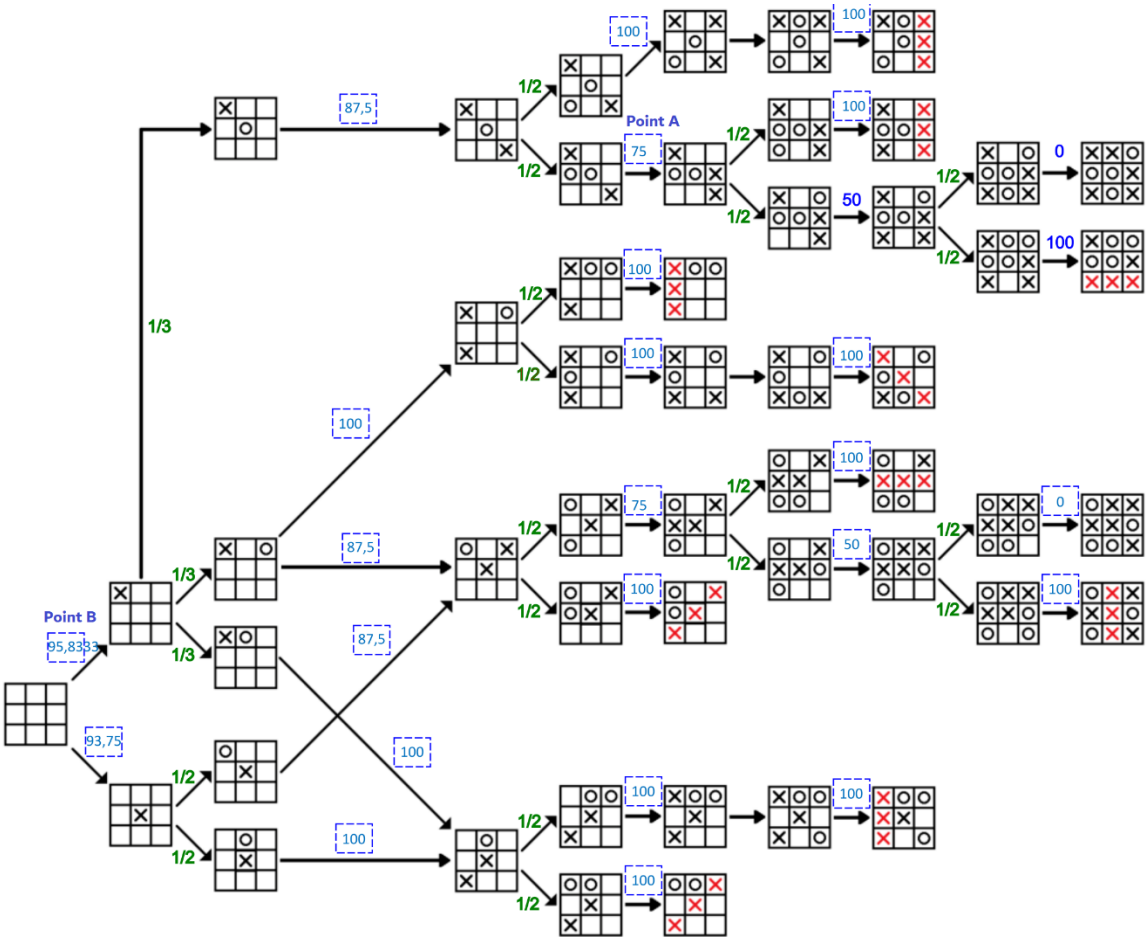
  (noter que comme la partie est finie, la partie $\gamma\, max_{a'}\, Q(s',a')$ est nulle)

- de même un deuxième cas conduit à la victoire :

  $$Q(s,a) = E[r + \gamma\, max_{a'}\, Q(s',a')] = E[r] = 100$$

- dans un troisième cas, on joue un coup tel que, ensuite, le deuxième joueur va jouer en bas avec probabilité ½ (récompense immédiate $r = 0$, nouvel état $s'$ tel que $max_{a'}\, Q(s',a') = 0$), et jouer en haut avec probabilité ½ ($r = 0$, nouvel état $s'$ tel que $max_{a'}\, Q(s',a') = 100$) ; on a donc (rappelons que $\gamma = 1$)

$$Q(s,a) = E[r + \gamma\, max_{a'}\, Q(s',a')] = 1/2 * (0 + 0) + 1/2 * (0 + 100) = 50$$




Explications supplémentaires pour deux exemples :

- Point A : après mon coup, le deuxième joueur jouera ensuite en bas à gauche avec probabilité ½ (récompense immédiate $r = 0$, nouvel état $s'$ tel que $max_{a'} Q(s', a') = 100$), et jouer en haut à droite avec probabilité ½ ($r = 0$, nouvel état $s'$ tel que $max_{a'} Q(s', a') = 50$) ; on a donc

$$Q(s, a) = E[r + \gamma \, max_{a'} Q(s', a')] = 1/2 * (0 + 100) + 1/2 * (0 + 50) = 75$$

  (mon coup vaut la moyenne entre 100 et 50, soit 75)

- Point B : après mon coup, le deuxième joueur jouera ensuite au centre avec probabilité ⅓ (récompense immédiate $r = 0$, nouvel état $s'$ tel que $max_{a'} Q(s', a') = 87.5$), en haut à droite avec probabilité ⅓ ($r = 0$, nouvel état $s'$ tel que $max_{a'} Q(s', a') = 100$ ; notez bien ici que dans le nouvel état $s'$ j'aurai le choix à mon tour entre deux coup, l'un valant 100 et l'autre 87.5, je choisirai bien sûr le maximum 100), et en haut avec probabilité ⅓ ($r = 0$, nouvel état $s'$ tel que $max_{a'} Q(s', a') = 100$)

$$Q(s, a) = E[r + \gamma \, max_{a'} Q(s', a')] = 1/3 * (0 + 87.5) + 1/3 * (0 + 100) + 1/3 * (0 + 100) = 95.8$$

  (mon coup vaut la moyenne entre 87.5, 100 et 100, soit 95.83)

Compléter les Q-valeurs limites à l'aide de l'équation de Bellman. Les emplacements à compléter sont indiqués par des rectangles bleus. Vous remarquerez qu'il faut commencer par remplir les valeurs les plus à droite.

On pourra arrondir les valeurs à l'unité la plus proche.

**Question 7 :**

Dans la configuration étudiée ici, et en supposant que la première joueuse choisit toujours l'action avec la Q-valeur la plus élevée, quelle est la probabilité de victoire pour la première joueuse (symbole X) ?
95.84% (c'est-à-dire la plus haute des 2 valeurs pour la 1ère action)

Avec les mêmes hypothèses, quelle est la probabilité d'un match nul ?
4.16% (1 – la réponse précédente)

# Exercice 6 : Étude de document

*Pour cet exercice, veuillez lire les textes et entourer/surligner la ou les bonne(s) réponse(s) ci-dessous.*

*Images obtenues en utilisant DALL-E 2*

DALL-E est un générateur d'images par IA conçu par OpenAI, connu pour avoir développé ChatGPT. Il permet de créer des visuels de toutes sortes à partir d'un prompt textuel. Son nom est un mot-valise évoquant à la fois le robot de Pixar WALL-E et le peintre Salvador Dalí. DALL-E est une version de 12 milliards de paramètres de GPT-3, entraînée pour générer des images à partir de descriptions textuelles, en utilisant un ensemble de données d'image annotées.

Q1. Qu'est-ce que DALL-E ?

a. Un modèle de traitement du langage naturel.
b. Un modèle d'IA générative pour la création d'images.
c. Un algorithme de reconnaissance d'objets.

Q2. Sur quoi DALL-E a-t-il été formé pour générer des images ?

a. Uniquement des descriptions textuelles.
b. Des paires texte-image.
c. Des images uniquement.

Q3. D'après ce qu'il est décrit au-dessus, est-ce que DALL-E ?

a. Est basée sur un apprentissage non-supervisé.
b. Est basée sur un apprentissage par renforcement.
c. Est basée sur un apprentissage supervisé.

Le fonctionnement de DALL-E repose sur une architecture de réseau de neurones appelée générateur d'images, qui prend en entrée une description textuelle et produit une image correspondante. Voici les étapes clés du processus :

- Pré-entraînement : DALL-E est initialement pré-entraîné sur un large ensemble de données contenant des paires texte-image. Ce pré-entraînement permet au modèle d'apprendre la structure sous-jacente des données visuelles et textuelles.
- Encodage de la description : Lorsqu'une description textuelle est fournie en entrée, le modèle utilise un encodeur pour convertir cette description en une représentation vectorielle, également appelée espace latent. Cette représentation capture les caractéristiques sémantiques de la description.

- Décodeur d'images : Le vecteur latent est ensuite fourni au décodeur d'images, qui génère une image correspondante. Ce décodeur est capable de transformer la représentation vectorielle en une image réaliste et cohérente en utilisant les connaissances acquises lors du pré-entraînement.
- Optimisation : Le modèle est optimisé pour minimiser la divergence entre l'image générée et les exemples d'entraînement réels. Cela permet d'affiner les paramètres du générateur afin d'améliorer la qualité des images générées.

Q4. Qu'est-ce que l'espace latent représente dans le contexte de DALL-E ?

a. L'espace des représentations textuelles.
b. La structure sous-jacente des données visuelles.
c. La dimension des images générées.

Q5. Comment le décodeur d'images utilise-t-il le vecteur latent pour générer une image correspondante ?

a. En minimisant la divergence.
b. En affinant les paramètres du générateur.
c. En utilisant les connaissances acquises lors du pré-entraînement.

Q6. Pourquoi DALL-E est-il soumis à une phase d'optimisation ?

a. Pour maximiser la divergence entre l'image générée et les exemples d'entraînement réels.
b. Pour minimiser la divergence entre l'image générée et les exemples d'entraînement réels.
c. Pour ajuster l'encodeur du modèle.

Afin de rendre DALL-E accessible à un large public, OpenAI a pris des mesures pour réduire les risques liés à ces modèles puissants de génération d'images. Des garde-fous ont été mis en place pour éviter que les images générées enfreignent la politique de contenu d'OpenAI. DALL-E est formé sur des centaines de millions d'images légendées provenant d'Internet, et OpenAI supprime et réévalue certaines de ces images pour influencer l'apprentissage du modèle. Cette stratégie vise notamment à filtrer les images violentes et sexuelles du jeu de données d'entraînement de DALL-E. Cette étape cruciale

a pour objectif d'empêcher le modèle d'apprendre à générer du contenu graphique ou explicite en réponse aux requêtes des utilisateurs, évitant ainsi la production involontaire d'images inappropriées. Cependant, cette démarche peut entraîner des biais potentiels, car les modèles formés sur des données filtrées ont tendance à amplifier certaines préférences, comme générer plus d'images mettant en scène des hommes et moins d'images représentant des femmes par rapport aux modèles formés sur l'ensemble de données original et non filtré.

Q7. Quelle est la principale raison pour laquelle OpenAI a mis en place des garde-fous lors de la mise à disposition de DALL-E au public ?

a. Pour empêcher les utilisateurs d'accéder aux fonctionnalités avancées de DALL-E.
b. Pour éviter que les images générées enfreignent la politique de contenu d'OpenAI.
c. Afin de limiter la création d'images susceptibles d'être utilisées à des fins non éthiques ou de causer du tort à autrui.

Q8. Quelle est la principale raison pour laquelle OpenAI supprime et réévalue certaines images du jeu de données d'entraînement de DALL-E ?

a. Pour réduire la taille du jeu de données.
b. Pour influencer délibérément l'apprentissage du modèle.
c. Pour accélérer le processus d'entraînement de DALL-E.

Sources :


- https://www.assemblyai.com/blog/how-dall-e-2-actually-works/#what-is-dall-e-
- https://openai.com/blog